

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2001, 10:52:58 ; Search time 16.02 seconds

(without alignments)
10.079 Million cell updates/sec

Title: US-09-288-719-1

Perfect score: 28

Sequence: 1 GGCGS 5

Scoring table: BIOSUM62

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	28	100.0	69 1 MCBM_ECOLI	P05834 escherichia
2	28	100.0	82 1 NIDM_MOUSE	P52503 mus musculus
3	28	100.0	104 1 HOI3_HOHI	Q23055 holotrichia
4	28	100.0	110 1 FUS1_HUMAN	O75896 homo sapien
5	28	100.0	114 1 ETS4_DROME	P29775 drosophila
6	28	100.0	118 1 CRYP_CRYPA	P18988 chenopodium
7	28	100.0	144 1 GRP1_CHERU	P37703 daucus caro
8	28	100.0	148 1 GRP9_DAUCA	Q10464 tritlicum ae
9	28	100.0	157 1 GRP1_WHEAT	Q03878 daucus caro
10	28	100.0	165 1 GRP1_ORYSA	P25074 oryza sativ
11	28	100.0	176 1 GRP7_ARATH	Q03250 arabidopsis
12	28	100.0	179 1 CBRFA_MAIZE	P25209 zea mays (m
13	28	100.0	182 1 KLEC3_BOVIN	P04261 bos taurus
14	28	100.0	183 1 OLEC_BRANA	P29526 brassica na
15	28	100.0	190 1 XYN_TRIHA	P48793 trichoderma
16	28	100.0	192 1 VC49_BPMD2	Q15672 homo sapien
17	28	100.0	202 1 TWST_HUMAN	P39694 bacillus su
18	28	100.0	205 1 CME1_MOUSE	P26687 mus musculi
19	28	100.0	206 1 TWST_MOUSE	P27484 nicotiana s
20	28	100.0	214 1 ERF2_NICOT	O80040 arabidopsis
21	28	100.0	222 1 ERF4_ARATH	O28250 canis famill
22	28	100.0	226 1 SPC1_CANFA	O94716 streptomyce
23	28	100.0	231 1 RNH_STRCO	P55007 rattus norv
24	28	100.0	235 1 BARL_RAT	O75190 homo sapien
25	28	100.0	241 1 HST2_HUMAN	P10495 phaseolus v
26	28	100.0	252 1 GRP1_PHAVU	O55834 synechocyst
27	28	100.0	254 1 EXB1_SYNY3	O52459 pseudomonas
28	28	100.0	259 1 NAHB_PSEPU	P09920 human cytom
29	28	100.0	261 1 USI3_HOMVA	P47802 homo sapien
30	28	100.0	265 1 CDX1_HUMAN	Q14681 homo sapien
31	28	100.0	265 1 Y176_HUMAN	P30794 zymomonas m
32	28	100.0	265 1 YPG1_ZYMMO	
33	28	100.0	265 1	

34	28	100.0	268 1 CDX1_MOUSE	P18111 mus musculu
35	28	100.0	268 1 EP34_HOMVA	P16768 human cytom
36	28	100.0	269 1 CH1E_MAIZE	P29023 zea mays (m
37	28	100.0	274 1 HAT9_ARATH	P46603 arabidopsis
38	28	100.0	276 1 KC2D_ARATH	O81275 arabidopsis
39	28	100.0	278 1 T2D7_DROME	Q27272 drosophila
40	28	100.0	280 1 CH1A_MAIZE	P29022 zea mays (m
41	28	100.0	283 1 KC2E_ARATH	O80507 arabidopsis
42	28	100.0	284 1 CDX4_HUMAN	O14627 homo sapien
43	28	100.0	288 1 CHIP_BETVU	P42820 beta vulgar
44	28	100.0	288 1 HME2_CHICK	O05917 gallus gall
45	28	100.0	297 1 SG54_DROME	Q00725 drosophila

ALIGNMENTS

RESULT 1
MCBA_ECOLI ID MCBM_ECOLI STANDARD; PRT; 69 AA.
AC P05834;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE BACTERIOCIN MICROICIN B17 PRECURSOR (MCB17).
GN MCBM.
OS Escherichia coli.
OG Plasmid IncFII pmcCB17.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88217867; PubMed-3329729;
RA Davagnino J., Herrero M., Furlong D., Moreno F., Koller R.;
RT "The DNA replication inhibitor microcin B17 is a
RL forty-three-amino-acid protein containing sixty percent glycine.";
RL Proteins 1:230-238(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-89123111; PubMed-2644225;
RA Genilloud O., Moreno F., Koller R.;
RT "DNA sequence, products, and transcriptional pattern of the genes
RT involved in production of the DNA replication inhibitor microcin
RT B17.";
RL J. Bacteriol. 171:1126-1135(1989).
RN [3]
RP SEQUENCE OF 1-14 FROM N.A.
RX MEDLINE-88216163; PubMed-2835580;
RA Conell N., Han Z., Moreno F., Koller R.;
RT "An E. coli promoter induced by the cessation of growth.";
RL Mol. Microbiol. 1:195-201(1987).
RN [4]
RP PARTIAL SEQUENCE OF 27-69.
RX MEDLINE-91122055; PubMed-1846808;
RA Vizan J.L., Hernandez-Chico C., del Castillo I., Moreno F.;
RT "Isolation and structure elucidation of the 43-peptide antibiotic
RL (In) Schneider C.H., Eberles A.N. (eds.).
RL Peptides 1992, pp.117-118, ESCOM Science Publishers, Leiden (1993).
RN [5]
RP FUNCTION.
RX MEDLINE-91122055; PubMed-1846808;
RA Vizan J.L., Hernandez-Chico C., del Castillo I., Moreno F.;
RT "The peptide antibiotic microcin B17 induces double-strand cleavage
RT of DNA mediated by E. coli DNA gyrase.";
RL EMBO J. 10:467-476(1991).
RN [6]
RP STRUCTURE BY NMR OF 1-26.
RX MEDLINE-98213789; PubMed-9545435;
RA Roy R.S., Kim S., Balleja J.D., Walsh C.T.;
RT "Role of the microcin B17 propeptide in substrate recognition:
RT solution structure and mutational analysis of McdA-26.";
RL Chem. Biol. 5:217-228(1998).
CC -!- FUNCTION: THIS GLYCINE-RICH PEPTIDE ANTIBIOTIC INHIBITS DNA

```
CC REPLICATION IN MANY ENTERIC BACTERIA, THAT LEADS TO INDUCTION OF
CC THE SOS REPAIR SYSTEM, MASSIVE DNA DEGRADATION AND CELL DEATH.
CC B17 INHIBITS TYPE II TOPISOMERASE BY TRAPPING AN ENZYME - DNA
CC CLEAVABLE COMPLEX.
CC -1- PPM: THE PROCESSED N-TERMINUS DOES NOT RESEMBLE A TYPICAL
CC SECRETION SIGNAL SEQUENCE.
CC -1- PPM: THE CYS RESIDUES AS WELL AS SOME GLY AND CYS ARE POST-
CC TRANSLATIONALLY MODIFIED. MODIFICATIONS INCLUDE THE FORMATION OF
CC FOUR THIAZOLE AND FOUR OXAZOLE RINGS THAT RESULT, RESPECTIVELY,
CC FROM THE CONDENSATION OF FOUR SERINE SIDE CHAINS WITH THE CARBONYL
CC GROUP OF THE PRECEDING AMINO ACID.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M15469; AAA24141.1; -
CC EMBL; M24253; AAA72741.1; -
CC EMBL; X06417; CAA29725.1; -
CC PIR; A25219; MIEC77.
CC PIR; A32058; A32058.
CC PDB; 2MLP; 22-JUL-98.
CC DNA replication inhibitor; Antibiotic; Bacteriocin; Plasmid;
CC 3D-structure.
CC PROPEP 1 26 BACTERIOCIN MICROCIN B17.
CC CHAIN 27 69 POLY-GLY.
CC DOMAIN 26 39
CC SEQUENCE 69 AA; 6013 MW; 0B1D159A832638A8 CRC64;
CC -----
Query Match 100.0%; Score 28; DB 1; Length 69;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGS 5
| | | | |
Db 36 GGGGS 40
-----
RESULT 2
NUMM MOUSE STANDARD; PRT; 82 AA.
ID NUMM MOUSE
AC P52503;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE 13 KDA-A SUBUNIT (EC 1.6.5.3)
DE (EC 1.6.99.3) (COMPLEX I-13KD-A) (CI-13KD-A) (FRAGMENT).
DE NDUS6 OR IPI3
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-CD-1;
RC MEDLINE-95331630; PubMed-7607554;
RA Watson J.D., Beckett-Jones B., Roy R.N., Green N.C., Flynn T.G.;
RT "Genomic sequence, structural organization and evolutionary
RT conservation of the 13.2-kda subunit of rat NADH:ubiquinone
RT oxidoreductase."
RL Gene 158:275-280(1995).
CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE. THIS IS A COMPONENT OF THE IRON-SULFUR (IP)
CC FRAGMENT OF THE ENZYME.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE; MATRIX SIDE.
CC -----
```

```
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L38438; AAB64010.1; -
CC MGD; MGI:107932; NDUS6.
CC Oxidoreductase; NAD; ubiquinone; Mitochondrion.
CC NON_TER 1
CC SEQUENCE 82 AA; 9330 MW; C923FFE245A9BD27 CRC64;
CC -----
Query Match 100.0%; Score 28; DB 1; Length 82;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGS 5
| | | | |
Db 50 GGGGS 54
-----
RESULT 3
HOLD3_HOLD1 STANDARD; PRT; 104 AA.
ID HOLD3_HOLD1
AC Q25055;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HOLOTRICIN 3 PRECURSOR.
DE HOLOTRICIN 3 PRECURSOR.
DE HOLOTRICIN 3 PRECURSOR.
DE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Scarabaeiformia; Scarabaeidae; Melolonthinae; Holoetricia.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-40.
RX TISSUE-LARVAL HEMOLYMPH;
RC MEDLINE-96073722; PubMed-8535393;
RA Lee S.Y., Moon H.J., Kurata S., Natori S., Lee B.L.;
RT "Purification and cDNA cloning of an antifungal protein from the
RT hemolymph of Holoetricia diomphalia larvae."
RL Biol. Pharm. Bull. 18:1049-1052(1995).
CC -1- FUNCTION: HOLOTRICIN 3 HAS ANTIFUNGAL ACTIVITY AGAINST
CC C. ALBICANS.
CC -1- SIMILARITY: TO TENECIN 3.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D13744; BAA02889.1; -
CC HSSP; P30129; ADPV.
CC Insect immunity; Antibiotic; Hemolymph; Fungicide; Signal; Repeat.
CC SIGNAL 1 20 HOLOTRICIN 3.
CC PEPTIDE 21 104 18 x 4 AA APPROXIMATE TANDEM REPEATS OF
CC DOMAIN 27 98 H-G-G-G.
CC REPEAT 27 30 1.
CC REPEAT 31 34 2.
CC REPEAT 35 38 3.
CC REPEAT 39 42 4.
CC REPEAT 43 46 5.
CC REPEAT 47 50 6.
CC REPEAT 51 54 7.
CC REPEAT 55 58 8.
CC REPEAT 59 62 9.
```

FT REPEAT 63 66 10.
 FT REPEAT 67 70 11.
 FT REPEAT 71 74 12.
 FT REPEAT 75 78 13.
 FT REPEAT 79 82 14.
 FT REPEAT 83 86 15.
 FT REPEAT 91 90 16.
 FT REPEAT 97 94 17.
 FT REPEAT 96 98 18.
 SO SEQUENCE 104 AA; 9026 MW; 2799D681BFDC725 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 28; DB 1; Length 104;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5
 |||||
 Db 69 GGGGS 73

RESULT 4
 FUS1_HUMAN STANDARD; PRT; 110 AA.
 AC 075896;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE FUS1 PROTEIN.
 GN FUS1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Latif F., Duh F.-M., Kuzmin I., Li H., Bader S., Sekido Y.,
 RA Minna J.D., Lerman M.I.;
 RT "FUS1, a highly conserved gene, is located in the smallest lung cancer
 RT region on 3p21.3";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- DISEASE: COULD BE THE PRODUCT OF A GENE INVOLVED IN LUNG CANCER.
 CC -1- SIMILARITY: BELONGS TO THE FUS1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL; AF055479; AAC35497.1;
 CC DR SEQUENCE 110 AA; 12074 MW; 9503BD10637C1504 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 28; DB 1; Length 110;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5
 |||||
 Db 19 GGGGS 23

RESULT 5
 ETS4_DROME STANDARD; PRT; 114 AA.
 ID ETS4_DROME
 AC P29775;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE DNA-BINDING PROTEIN D-ETS-4 (FRAGMENT).
 GN ETS98B OR ETS-4.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-92249640; PubMed-1577186;
 RA Chen T., Buntling M., Karim F.D., Thummel G.S.;
 RT "Isolation and characterization of five Drosophila genes that encode
 RT an ets-related DNA binding domain";
 RL Dev. Biol. 151:176-191(1992).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL; M88474; AAA28451.1;
 CC DR PIR; S28821; S28821.
 CC DR HSSP; Q01543; IFLI.
 CC DR FLYBASE; FBgn0005659; ETS98B.
 CC DR INTERPRO; IPR000418;
 CC DR PFAM; PF00178; Ets; 1.
 CC DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
 CC DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
 CC DR PROSITE; PS50061; ETS_DOMAIN_3; 1.
 CC KW DNA-binding; Nuclear protein.
 CC FT NON_TER 1
 CC FT DNA_BIND 26 108 ETS-DOMAIN.
 CC FT NON_TER 114 114
 CC SEQUENCE 114 AA; 13051 MW; 6D7B46CE006A9743 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 28; DB 1; Length 114;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5
 |||||
 Db 20 GGGGS 24

RESULT 6
 CRYP_CRYPA STANDARD; PRT; 118 AA.
 ID CRYP_CRYPA
 AC P52753;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE CRYPARIN PRECURSOR.
 CC CRYP.
 CC Cryphonectria parasitica (chestnut blight fungus) (Endothia
 CC parasitica).
 CC Eukaryota; Fungi; Ascomycota; Diaportheales; Valsaceae; Cryphonectria.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-73.
 RC STRAIN-155/2;
 RX MEDLINE-94156182; PubMed-8112589;
 RA Zhang L., Villalón D., Sun Y., Kazmierczak P., van Alfen N.K.;
 RT "Virus-associated down-regulation of the gene encoding cryptarin, an
 RT abundant cell-surface protein from the chestnut blight fungus,
 RT Cryphonectria parasitica";
 RL Gene 139:59-64(1994).
 CC -1- FUNCTION: CONTRIBUTES TO SURFACE HYDROPHOBICITY, WHICH IS
 CC IMPORTANT FOR PROCESSES SUCH AS ASSOCIATION OF HYPAE IN
 CC REPRODUCTIVE STRUCTURES, DISPERSAL OF AERIAL SPORES AND ADHESION
 CC OF PATHOGENS TO HOST STRUCTURES. PRODUCED ABUNDANTLY, EXCEPT IN
 CC THE DS-RNA VIRUS-INFECTED STRAINS, WHERE THE EXPRESSION IS MUCH

```

CC REDUCED.
CC -1- SUBCELLULAR LOCATION: CELL WALL OF AERIAL HYPHAE AND SPOULATION
CC STRUCTURES. ABUNDANTLY SECRETED.
CC -1- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED ON DAY 2 AND 3 AFTER
CC INOCULATION. A TIME WHEN THE FUNGUS IS IN A RAPID PHASE OF GROWTH.
CC AFTER A STATIONARY PHASE ON DAY 4, THE EXPRESSION DECREASES.
CC -1- SIMILARITY: BELONGS TO THE CERATO-ULMIN HYDROPHOBIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L09559; AAA19638.1; -
FT SIGNAL 1 22
FT CHAIN 23 118 CRYPARIN.
FT DOMAIN 23 32 POLY-GLY.
FT REPEAT 29 42 7 X 2 AA TANDEM REPEAT OF S-G.
FT REPEAT 29 30 1.
FT REPEAT 31 32 2.
FT REPEAT 33 34 3.
FT REPEAT 35 36 4.
FT REPEAT 37 38 5.
FT REPEAT 39 40 6.
FT REPEAT 41 42 7.
SQ SEQUENCE 118 AA; 11387 MW; F7C7CCEA57D06A5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 28; DB 1; Length 118;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|11111
DB 25 GGGGS 29

RESULT 7
GRPL_CHERU STANDARD; PRT; 144 AA.
AC P11898;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-1984 (Rel. 30, Last annotation update)
DE GLYCINE-RICH PROTEIN HCl.
OS Chenopodium rubrum (Red goosefoot) (Pigweed).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
OC Caryophyllales; Chenopodiales; Chenopodium.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89240041; PubMed-2717413;
RA Kaldenhoff R., Richter G.;
RT "Sequence of cDNA for a novel light-induced glycine-rich protein.";
RL Nucleic Acids Res. 17:2853-2853(1989).
CC -1- INDUCTION: BY LIGHT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14067; CAA32230.1; -
CC PIR; S04069; S04069.
KW Repeat; Transmembrane.
FT TRANSMEM 5 25 POTENTIAL.

```

```

FT DOMAIN 37 113 11 X 6 AA TANDEM REPEATS OF G-Y-[NH]-N-G
FT REPEAT 37 42 -G.
FT REPEAT 43 48 1.
FT REPEAT 50 55 2.
FT REPEAT 56 61 3.
FT REPEAT 63 68 4.
FT REPEAT 69 74 5.
FT REPEAT 76 81 6.
FT REPEAT 82 87 7.
FT REPEAT 89 94 8.
FT REPEAT 102 107 9.
FT REPEAT 108 113 10.
FT REPEAT 113 113 11.
SQ SEQUENCE 144 AA; 14137 MW; 5B4D62DA4A61621B0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 28; DB 1; Length 144;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|11111
DB 111 GGGGS 115

RESULT 8
GRP9_DAUCA STANDARD; PRT; 144 AA.
AC P37703;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE GLYCINE-RICH PROTEIN DC9.1.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae;
OC eunsterids II; Apiales; Apiaceae; Daucus.
RN [1]
RP SEQUENCE FROM N.A.
RA Aleith F., Richter G.;
RT "Gene expression during induction of somatic embryogenesis in carrot
RT cell suspensions.";
RL Planta 183:17-24(1990).
DR PIR; S35716; S35716.
KW Repeat; Transmembrane.
FT TRANSMEM 5 25
FT DOMAIN 37 113
FT REPEAT 37 42 -G.
FT REPEAT 43 48 1.
FT REPEAT 48 55 2.
FT REPEAT 50 55 3.
FT REPEAT 56 61 4.
FT REPEAT 63 68 5.
FT REPEAT 69 74 6.
FT REPEAT 76 81 7.
FT REPEAT 82 87 8.
FT REPEAT 89 94 9.
FT REPEAT 102 107 10.
FT REPEAT 108 113 11.
SQ SEQUENCE 144 AA; 14111 MW; 5B4D62CFBCA791B0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 28; DB 1; Length 144;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
|11111
DB 111 GGGGS 115

RESULT 9
PUB_WHEAT

```

ID PUB.WHEAT STANDARD; PRT; 148 AA.
 AC Q10464;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PROINDOLINE-B PRECURSOR.
 OS Trilicium aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Triticum.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. CAPITOLE; TISSUE-SEED;
 RX MEDLINE-94272013; PubMed-7516201;
 RA Gautier M.-F., Aleman M.-F., Guirao A., Marion D., Joudrier P.;
 RT "Triticum aestivum proindolines, two basic cysteine-rich seed
 RT proteins: cDNA sequence analysis and developmental gene expression.";
 RT Plant Mol. Biol. 25:43-57(1994).
 RN [2]
 RP SEQUENCE OF 30-148.
 RA Blochet J.E., Kaboulou A., Compoin J.P., Marion D.;
 RL (In) Bushuk W., Tkachuk R. (eds.);
 RL Gluten proteins, pp.314-325, American Association of Cereal Chemists,
 RL St. Paul MI (1991).
 CC -1- FUNCTION: ACTS AS A MEMBRANOTOXIN, PROBABLY THROUGH ITS
 CC ANTIBACTERIAL AND ANTIFUNGAL ACTIVITIES, CONTRIBUTING TO THE
 CC DEFENSE MECHANISM OF THE PLANT AGAINST PREDATORS.
 CC -1- PTM: FIVE DISULFIDE BONDS ARE PRESENT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X69912; CAA49537.1; -
 DR HSSP: P07597; 1LTP
 DR INTERPRO: IPR001768; -
 DR PFAM: PF00234; tryp_alpha_amy1; 1.
 KM Membrane; Toxin; Antibiotic; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT PROPEP 20 29
 FT CHAIN 30 148
 FT DOMAIN 68 73 PROINDOLINE-B.
 FT SEQUENCE 148 AA; 16792 MW; 327904B4EBEC2C16 CRC64;
 SQ
 Query Match 100.0%; Score 28; DB 1; Length 148;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GGGGS 5
 DB 33 GGGGS 37
 RESULT 10
 GRP DAUCA STANDARD; PRT; 157 AA.
 AC 003878;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE GLYCINE-RICH RNA-BINDING PROTEIN.
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae;
 OC euasterids II; Apiales; Apiaceae; Daucus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. QUEEN ANNE'S LACE;
 RA Sturm A.;

RT "A wound-inducible glycine-rich protein from Daucus carota with
 RT homology to single-stranded nucleic acid binding proteins.";
 RL Plant Physiol. 99:1689-1692(1992).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE BIOSYNTHESIS AND PROCESSING OF
 CC HETEROGENEOUS NUCLEAR RNA AND IN THE MATURATION OF SPECIFIC MRNAS
 CC IN RESPONSE TO WOUNDING.
 CC -1- INDUCTION: IN RESPONSE TO STRESS BY WOUNDING.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RNM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X58146; CAA41152.1; -
 DR PIR: S14857; S14857.
 DR HSSP: P09651; 1UP1.
 DR INTERPRO: IPR000504; -
 DR PFAM: PF00076; rrm; 1.
 DR PROSITE: PS50102; RRM; 1.
 DR PROSITE: PS00030; RRM_RNP_1; 1.
 KW RNA-binding.
 FT DOMAIN 8 13 RNA-BINDING (RNP2) (BY SIMILARITY).
 FT DOMAIN 47 54 RNA-BINDING (RNP1) (BY SIMILARITY).
 FT DOMAIN 86 154 GLY-RICH
 SQ SEQUENCE 157 AA; 15718 MW; 73FBD644F51CB633 CRC64;
 OY 1 GGGGS 5
 DB 151 GGGGS 155
 Query Match 100.0%; Score 28; DB 1; Length 157;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GGGGS 5
 DB 151 GGGGS 155
 RESULT 11
 GRP1-ORFSA STANDARD; PRT; 165 AA.
 AC P25074;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN 1 PRECURSOR.
 GN GRP-1.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. INDICA-IR36;
 RX MEDLINE-91370862; PubMed-1716496;
 RA Lei M., Wu R.;
 RT "A novel glycine-rich cell wall protein gene in rice.";
 RL Plant Mol. Biol. 16:187-198(1991).
 CC -1- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: CELL WALL (POTENTIAL).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X53596; CAA37665.1; -
 DR PIR: S13385; KNRZG1.
 DR HSSP: P30129; 4DPV.

KW Cell wall; Structural protein; Repeat; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 165 GLYCINE-RICH CELL WALL STRUCTURAL
 FT DOMAIN 31 159 PROTEIN 1.
 FT REPEAT 56 62 GLY-RICH.
 FT REPEAT 93 99 R2 (TYR-RICH).
 FT REPEAT 132 138 R2 (TYR-RICH).
 FT REPEAT 132 138 R2 (TYR-RICH).
 SO SEQUENCE 165 AA; 13536 MW; E36CE31C3650AC9A CRC64;

Query Match 100.0%; Score 28; DB 1; Length 165;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 Db 121 GGGGS 125

RESULT 12
 GRP7_ARATH STANDARD; PRT; 176 AA.
 ID GRP7_ARATH STANDARD; PRT; 176 AA.
 AC 003250:
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE GLYCINE-RICH RNA-BINDING PROTEIN 7.
 GN GRP7.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 CC Brassicales; Brassicaceae; Arabidopsids.
 RN [1].
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA; TISSUE=LEAF;
 RX MEDLINE-93192529; PubMed-8448367;
 RA van Nocker S.; Verstra R.D.;
 RT "Two cDNAs from Arabidopsis thaliana encode putative RNA binding
 proteins containing glycine-rich domains.";
 RL Plant Mol. Biol. 21:695-699(1993).
 CC -1- FUNCTION: POSSIBLY HAS A ROLE IN RNA TRANSCRIPTION OR
 PROCESSING DURING STRESS.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: Z14987; CAA7871.1; -
 CC EMBL: L00648; AAA32853.1; -
 CC PIR: S30147; S30147.
 CC HSSP: P09651; IUP1.
 CC INTERPRO: IPR000504; -
 CC PIR: P000076; IUP1.
 CC PROSITE: PS50102; RRM_1.
 CC PROSITE: PS00030; RRM_RNP_1.
 CC RNA-binding.
 FT DOMAIN 10 15 RNA-BINDING (RNP2) (BY SIMILARITY).
 FT DOMAIN 49 56 RNA-BINDING (RNP1) (BY SIMILARITY).
 FT DOMAIN 88 176 GLY-RICH
 SO SEQUENCE 176 AA; 16890 MW; 3E1025477F9CF4C4 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 176;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5

Db 113 GGGGS 117

RESULT 13
 CBFA_MAIZE STANDARD; PRT; 179 AA.
 ID CBFA_MAIZE STANDARD; PRT; 179 AA.
 AC P25209:
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE CCAAT-BINDING TRANSCRIPTION FACTOR SUBUNIT A (CBF-A) (NF-Y PROTEIN
 DE CHAIN B) (NF-YB) (CAAT-BOX DNA BINDING PROTEIN SUBUNIT B).
 GN NFY2.
 OS Zea mays (Maize).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; liliopsida; Poales; Poaceae; Zea.
 RN [1].
 RP SEQUENCE FROM N.A.
 RC MEDLINE-92195809; PubMed-1549471;
 RA Li X.-Y.; Mantovani R.; Hooft van Hujsduijn R.; Andre I.;
 RA Benoist C.; Mathis D.;
 RT "Evolutionary variation of the CCAAT-binding transcription factor
 NF-Y.";
 RL Nucleic Acids Res. 20:1087-1091(1992).
 CC -1- FUNCTION: STIMULATES THE TRANSCRIPTION OF VARIOUS GENES BY
 RECOGNIZING AND BINDING TO A CCAAT MOTIF IN PROMOTERS. FOR
 CC EXAMPLE IN TYPE I COLLAGEN, ALBUMIN AND BETA-ACTIN GENES.
 CC -1- SUBUNIT: HETEROMERIC TRANSCRIPTION FACTOR COMPOSED OF TWO
 CC COMPONENTS, A AND B, THAT ARE BOTH NEEDED FOR DNA BINDING.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- DOMAIN: CAN BE DIVIDED INTO THREE DOMAINS: THE WEAKLY CONSERVED A
 CC DOMAIN, THE HIGHLY CONSERVED B DOMAIN THOUGHT TO BE INVOLVED IN
 CC SUBUNIT INTERACTION AND DNA BINDING, AND THE GLU-RICH C DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE CBF-A SUBUNIT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X59714; CAA42234.1; -
 CC HSSP: P19267; IBEF.
 CC MAIZEEDB: 69282; -
 CC INTERPRO: IPR000947; -
 CC PIR: P00808; CBF_NFYB_HMF_1.
 CC PROSITE: PR00615; CCAATSUBUNITA.
 CC PROSITE: PR00685; CBFA_NFYB_1.
 CC TRANSCRIPTION regulation; DNA-binding; Activator; Nuclear protein.
 FT DOMAIN 1 29 A DOMAIN.
 FT DOMAIN 30 119 B DOMAIN.
 FT DOMAIN 120 179 C DOMAIN.
 FT DNA_BIND 36 42 BY SIMILARITY.
 SO SEQUENCE 179 AA; 18995 MW; 30621316CEA69454 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 179;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
 Db 10 GGGGS 14

RESULT 14
 K2C3_BOVIN STANDARD; PRT; 182 AA.
 ID K2C3_BOVIN STANDARD; PRT; 182 AA.
 AC P04261;
 DT 20-MAR-1987 (Rel. 04, Created)

DT 20-MAR-1987 (Rel. 04, last sequence update)
 DT 01-JUN-1994 (Rel. 29, last annotation update)
 DE KERATIN, TYPE II CYTOSKELETAL 60 KDA, COMPONENT III (FRAGMENT).
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-85128114; PubMed-6084625.
 RA Jorcano J.L., Franz J.K., Franke W.W.;
 RT "Amino acid sequence diversity between bovine epidermal cyokeratin
 RT polypeptides of the basic (type II) subfamily as determined from cDNA
 RT clones".
 RL Differentiation 28:155-163 (1984).
 CC -1- SUBUNIT: HETEROTRIMER OF TWO TYPE I AND TWO TYPE II KERATINS.
 CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
 CC MICROFIBRILLAR KERATIN, I (ACIDIC) AND II (NEUTRAL TO BASIC)
 CC (40-55 AND 56-70 KILODALTONS, RESPECTIVELY).
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; K03535; AAA30602.1; -
 DR PIR; A02947; A02947.
 DR INTERPRO: IPR001664; -
 DR PFAM: PF00038; filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Heptad repeat pattern; Keratin.
 FT NON_TER 1 1 ROD.
 FT DOMAIN <1 63 TAIL.
 FT DOMAIN 63 182 TAIL.
 FT DOMAIN <1 63 COIL 2.
 FT SITE 1 1 STUTTER.
 SQ SEQUENCE 182 AA; 17560 MW; C263102147632F66 CRC64;

 QY 1 GGGGS 5
 Db 119 GGGGS 123

 Query Match 100.0%; Score 28; DB 1; Length 182;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RL Plant J. 3:629-636(1993).
 CC -1- FUNCTION: MAY HAVE A STRUCTURAL ROLE TO STABILIZE THE LIPID BODY
 CC DURING DESTICATION OF THE SEED BY PREVENTING COALESCENCE OF THE
 CC OIL. PROBABLY INTERACTS WITH BOTH LIPID AND PHOSPHOLIPID MOETIES
 CC OF LIPID BODIES. MAY ALSO PROVIDE RECOGNITION SIGNALS FOR SPECIFIC
 CC LIPASE ANCHORAGE IN LIPOLYSIS DURING SEEDLING GROWTH.
 CC -1- SUBCELLULAR LOCATION: SURFACE OF OIL BODIES. OLEOSINS EXIST AT A
 CC MONOLAYER LIPID/WATER INTERFACE.
 CC -1- TISSUE SPECIFICITY: SPECIFIC TO THE MALE GAMETOPHYTE.
 CC -1- SIMILARITY: BELONGS TO THE OLEOSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X67142; CAA47623.1; -
 DR PIR; S24960; S24960.
 DR HSSP; P04002; IMFA.
 DR INTERPRO: IPR000136; -
 DR PROSITE: PS00811; OLEOSINS; 1.
 KW Seed; Oil body; Multigene family.
 FT NON_TER 1 1
 FT DOMAIN <1 23 POLAR.
 FT DOMAIN 24 95 HYDROPHOBIC.
 SQ SEQUENCE 183 AA; 18149 MW; 198A5D3B6DF3045A CRC64;

 QY 1 GGGGS 5
 Db 123 GGGGS 127

 Query Match 100.0%; Score 28; DB 1; Length 183;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: March 15, 2001, 11:11:10
 Job time: 1092 sec

